

Wherefore, what is claimed is:

1. A computer-implemented face detection process for detecting a person's face in an input image and identifying a face pose range into which the face pose exhibited by the detected face falls, comprising using a computer to perform the following process actions:

creating a database comprising a plurality of training feature characterizations, each of which characterizes the face of a person at a known face pose or a non-face;

training a plurality of detectors arranged in a pyramidal architecture to determine whether a portion of an input image depicts a person's face having a face pose falling within a face pose range associated with one of the detectors using the training feature characterizations; and wherein

said detectors using a greater number of feature characterizations are arranged at the bottom of the pyramid, and wherein

said detectors arranged to detect finer ranges of face pose are arranged at the bottom of the pyramid;

inputting a portion of an input image into the plurality of detectors arranged in a pyramid architecture; and

interpreting the output of the plurality of detectors to determine whether the portion of the input image contains a face and if so to identify the pose associated with each detected face.

2. The process of Claim 1 wherein each of said detectors of said plurality of detectors comprises at least one classifier, said at least one classifier employing a unique feature that characterizes a face with poses within the range of the detector, each classifier of the detector determined by employing a statistical process to identify the classifier or classifiers that best indicates that the portion of the input image under consideration is a face in the pose range associated with the detector.

3. The process of Claim 1 wherein the process action of creating a database comprises the actions of:

capturing training images of faces of a plurality of people at a variety of face poses; and

preprocessing said images to prepare them for input into the plurality of detectors arranged in a pyramid architecture.

4. The process of Claim 3 wherein the process action of preprocessing the images to prepare them for input into the detector-pyramid comprises the actions of:

normalizing each training image by resizing it to a prescribed scale if not already at the prescribed scale and adjusting the region so that the eye locations of the depicted subject fall within a prescribed area;

cropping each training image to eliminate unneeded portions not specifically depicting part of the face of the subject;

categorizing the normalized and cropped training images according to their face pose by grouping the images into a set of pose ranges.

5. The process of Claim 1 wherein each detector is constructed based on one or more weak classifiers.

6. The process of Claim 5 wherein each classifier performs face/non-face classification using a different single feature.

7. The process of Claim 1 wherein each detector can be one of:

a single face/non-face classifier; and

a cascade of face/non-face classifiers.

8. The process of Claim 1 wherein said detectors at the bottom of the pyramid using a greater number of feature characterizations and arranged to detect finer ranges of face pose are more complex.

9. The process of Claim 1 wherein training the detectors comprises one of:

using a Gaussian model;  
using a small set of simple image features and a neural network;  
using a boosting algorithm; and  
using a support vector machine.

10. The process of Claim 1 wherein training the detectors comprises the following process actions:

designing a set of simple features;  
selecting a subset of the set of simple features;  
training a set of weak classifiers using said subset of features;  
constructing a strong classifier from a linear combination of weak classifiers;

at each level of the pyramid, partitioning the full range of face poses into a number of sub-ranges and training the same number of detectors for face detection in each partition, each detector specialized for a certain pose sub-range; and  
composing the detector-pyramid of several levels from the coarsest view partition at the top to the finest view partition at the bottom.

11. The process of Claim 1 wherein each detector is designed to detect one face pose range associated with that detector.

12. The process of Claim 11 wherein the face pose range of a detector may partially overlap the face pose range associated with another detector.

13. The process of Claim 1 wherein detectors only on one side of the detector pyramid are trained, and detectors on the other side of the pyramid are deemed to be the mirrors of the trained side.

14. The process of Claim 13 wherein the training time of training the detectors of the detector pyramid is cut in half because only the detectors on one side of the detector pyramid are trained.

15. The process action of Claim 1, wherein the process action of inputting a region from the input image comprises the action of:  
partitioning said input image into sub-windows.

16. The process action of Claim 15 wherein said partitioning of said input image into sub-windows comprises moving a search window of a prescribed size across the input image and prior to each shift extracting the pixels contained within the search window to create an input image region.

17. The process of Claim 16, further comprising the process actions of:  
employing a search window having a size approximately equal to the size of the smallest face it is anticipated will be depicted in the input image and which it is desired to detect;

after regions from every part of the input image it is desired to screen for faces have been extracted, reducing the size of the input image by a prescribed scale increment;

progressively shifting the search window across the reduced input image and prior to each shift extracting the pixels contained within the search window to create an input image region; and

repeating the reducing and shifting process actions until a prescribed reduction limit is reached.

18. The process of Claim 16, wherein the search window size corresponds to the size of the training images.

19. The process of Claim 16, wherein the search window size is the size of the smallest detectable face anticipated to be found in the input image.

20. The process of Claim 19, wherein the search window size is 20 by 20 pixels.

21. The process of Claim 16, wherein the initial search window size is increased by a scale factor in a step-wise fashion all the way up to the input image size; and after each increase in scale partitioning the input image with the search sub-window size.

22. The process of Claim 16 wherein the original sub-window size matches the entire image and this sub-window is then scaled down on an incremental basis.

23. The process of Claim 1 wherein the detector pyramid architecture comprises three detector layers and wherein

said first detector layer comprises a single full-view detector responsible for the full range of -90 to 90 degrees of face pose, with 0 degrees being frontal view;

said second detector layer comprises a first, second and third detector, said first detector being capable of detecting face pose ranges of -90 to 40 degrees, said second detector being capable of detecting face pose ranges of -30 to 30 degrees, and said third detector being capable of detecting face pose range of 40 to 90 degrees;

said third detector layer comprising nine detectors, capable of detecting face pose ranges of -90 to -80 degrees, -70 to -60 degrees, -50 to -40 degrees, -30 to -20 degrees, -10 to 10 degrees, 20 to 30 degrees, 40 to 50 degrees, 60 to 70 degrees, and 80 to 90 degrees, respectively.

24. The process of Claim 1 wherein the process action of inputting a portion of an input image into the plurality of detectors arranged in a pyramid architecture comprises:

inputting a portion of the input image into a first detector layer;  
if the portion of the input image is rejected by the detector at the top layer,  
it is classified as a non-face region it is not processed by detectors in later detector layers;  
if the portion of the input image is processed by the detectors in the first  
5 detector layer, it is processed by the second layer, if a detector in the second layer  
classifies the input image portion as a non-face region it is not processed by detectors in  
the third layer;  
if the portion of the input image is processed by the detectors in the second  
10 detector layer, it is processed by the third detector layer, which classifies the input image  
region into a face pose range corresponding to a detector trained to detect a given face  
pose range.

25. The process of Claim 1 wherein inputting a portion of an input image into  
the plurality of detectors arranged in a pyramid architecture further comprises  
15 arbitrating between two or more detectors that detect a face in the same  
detector layer to determine if the detections represent two different faces or two  
detections of one face.

26. The process of Claim 25 wherein arbitrating between two or more  
20 detectors further comprises  
determining if the detections by each of the two or more detectors overlap;  
specifying that if the detections by each of the two or more detectors do  
not overlap then arbitration is not necessary and each face detection is determined to be a  
separate face;  
25 combining the output of some of the detector view ranges into one class by  
creating new classes of view ranges from the various pose range detectors at the detector  
pyramid's outputs;  
arbitrating between the new classes of view ranges to categorize each  
30 overlapping detection into one of the new classes of view ranges.

27. The process of Claim 26 wherein the arbitrating between the new classes of view ranges comprises using Rowley's heuristic method.

28. The process of Claim 27 wherein arbitrating between the new classes of view ranges comprises

determining whether a face detection at any detector is identified as a frontal face;

if said face detection is determined to be a frontal face then all other face locations detected by profile or half profile detectors that are overlapping in the input image are determined to be errors and are eliminated, and the face detection that is determined to be a frontal face is determined to be a single frontal face;

if the face detection is not identified as a frontal face, determining whether the given location is identified as a half frontal face;

if the location is identified as a half profile face then all other locations detected by profile face detectors are eliminated and the particular location is determined to be a half profile face; and

if the location is not a non-face, nor a frontal face, nor a half profile face, then the location is determined to be a profile face.

29. A system for detecting an object in an input image, the system comprising:  
a general purpose computing device; and  
a computer program comprising program modules executable by the computing device, wherein the computing device is directed by the program modules of the computer program to

create a database comprising a plurality of training feature characterizations, each of which characterizes an object being sought at a known orientation or a an object not being sought;

train a plurality of detectors arranged in a pyramidal architecture to determine whether a portion of an input image depicts an object being sought having an

orientation within an orientation associated with one of the detectors using the training feature characterizations; and wherein

said detectors using a greater number of feature characterizations are arranged at the bottom of the pyramid, and wherein

5 said detectors arranged to detect finer ranges of object orientation are arranged at the bottom of the pyramid;

input a portion of an input image into the plurality of detectors arranged in a pyramid architecture; and

10 interpret the output of the plurality of detectors to determine whether the portion of the input image contains an object being sought and if so to identify the orientation associated with each detected object being sought.

30. A computer-readable medium having computer-executable instructions for detecting a person's face in an input image, said computer executable instructions comprising:

15 creating a database comprising a plurality of training feature characterizations, each of which characterizes the face of a person at a known face pose or a non-face;

20 training a plurality of detectors arranged in a pyramidal architecture to determine whether a portion of an input image depicts a person's face having a face pose falling within a face pose range associated with one of the detectors using the training feature characterizations, said plurality of detectors when trained being capable of determining whether a portion of an input image depicts a person's face; and wherein

25 said detectors using a greater number of feature characterizations are arranged at the bottom of the pyramid, and wherein

said detectors arranged to detect finer ranges of face pose are arranged at the bottom of the pyramid.



31. A face detection system for detecting a person's face depicted in an input image and identifying a face pose range, among a set of pose ranges, into which the pose associated with the detected face falls, comprising:

5 a database comprising a plurality of training feature characterizations, each of which characterizes the face of a person at a known face pose or a non-face;

a plurality of detectors arranged in a pyramidal architecture to determine whether a portion of an input image depicts a person's face having a face pose falling within a face pose range associated with one of the detectors using said training feature characterizations; and wherein

10 said detectors using a greater number of feature characterizations are arranged at the bottom of the pyramid, and wherein

said detectors arranged to detect finer ranges of face pose are arranged at the bottom of the pyramid.